

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-099566

(43)Date of publication of application : 05.04.2002

(51)Int.Cl.

G06F 17/30

(21)Application number : 2000-291799

(71)Applicant : TOSHIBA CORP

(22)Date of filing : 26.09.2000

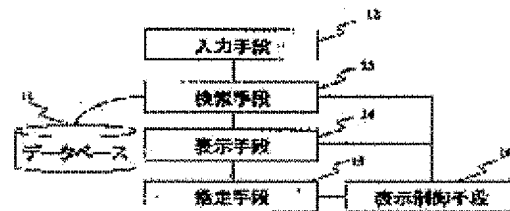
(72)Inventor : UDA AKIHIRO
ISHITANI YASUTO
KUBOTA HIROAKI

(54) METHOD AND DEVICE FOR INFORMATION RETRIEVAL

(57)Abstract:

PROBLEM TO BE SOLVED: To overcome the problem of a conventional method for information retrieval such that it takes much time and labor to narrow down objective information because the objective information has to be found while watching the contents (text information of regular text) by displaying information extracted by the first retrieval.

SOLUTION: In the present invention, text data including a retrieval character string are retrieved from a storage means by inputting the retrieval character string. Then, when desired text data are specified from a retrieval list of the retrieved text data, the number of the retrieval character strings included in the designated text data is displayed together with the contents of the retrieval character string.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2002-99566
(P2002-99566A)

(43) 公開日 平成14年4月5日 (2002.4.5)

(51) Int.Cl.⁷
G 0 6 F 17/30

識別記号
3 2 0
1 7 0
3 8 0

F I
G 0 6 F 17/30

テームコード* (参考)

3 2 0 D 5 B 0 7 5
1 7 0 A
3 8 0 E

審査請求 未請求 請求項の数13 O L (全 8 頁)

(21) 出願番号 特願2000-291799(P2000-291799)

(22) 出願日 平成12年9月26日 (2000.9.26)

(71) 出願人 000003078

株式会社東芝
東京都港区芝浦一丁目1番1号

(72) 発明者 宇田 明弘

神奈川県川崎市幸区小向東芝町1番地 株
式会社東芝研究開発センター内

(72) 発明者 石谷 康人

神奈川県川崎市幸区小向東芝町1番地 株
式会社東芝研究開発センター内

(74) 代理人 100083161

弁理士 外川 英明

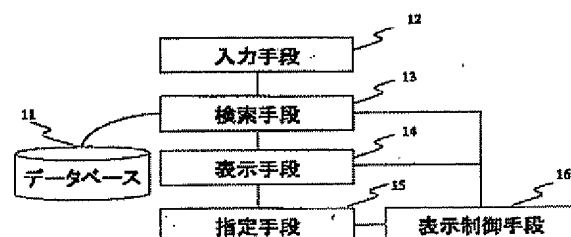
最終頁に続く

(54) 【発明の名称】 情報検索方法および情報検索装置

(57) 【要約】

【課題】従来の情報検索方法では、一次検索で抽出された情報を表示してその内容（本文のテキスト情報）を見ながら、目的とする情報を見つけていかなければならなかったため、目的とする情報を絞り込むことに非常に手間を要していた。

【解決手段】本発明においては、検索文字列を入力することで、この検索文字列を含むテキストデータを記憶手段から検索するときに、検索されたテキストデータの検索一覧リストから所望のテキストデータが指定されたとき、この指定されたテキストデータに含まれている検索文字列の個数を検索文字列の内容とともに表示することを特徴とする。



【特許請求の範囲】

【請求項 1】 検索対象となる複数のテキストデータが記憶されている記憶手段から所望のテキストデータを検索する情報検索方法において、

前記記憶手段に記憶されている複数のテキストデータから所望のテキストデータを検索するために検索文字列を入力し、

この入力された検索文字列を含むテキストデータを前記記憶手段から検索し、

この検索されたテキストデータの検索一覧リストを表示し、

この表示された検索一覧リストの中から所望のテキストデータを指定し、

この指定されたテキストデータに含まれている検索文字列の個数を検索文字列の内容とともに表示することを特徴とした情報検索方法。

【請求項 2】 前記検索文字列の内容とこの検索文字列が含まれている個数の表示する際に、前記検索一覧リストが表示されている領域に重ねて表示することを特徴とする請求項 1 記載の情報検索方法。

【請求項 3】 前記検索文字列の内容とこの検索文字列が含まれている個数の表示する際に、前記検索一覧リストが表示されている領域とは別の領域に表示することを特徴する請求項 1 記載の情報検索方法。

【請求項 4】 検索対象となる複数のテキストデータが記憶されている記憶手段から所望のテキストデータを検索する情報検索方法において、

前記記憶手段に記憶されている複数のテキストデータから所望のテキストデータを検索するために第 1 の検索文字列を入力し、

この入力された第 1 の検索文字列を拡張処理して第 2 の検索文字列を作成し、

前記第 1 の検索文字列及びこの第 1 の検索文字列を拡張処理した前記第 2 の検索文字列を含むテキストデータを前記記憶手段から検索し、

この検索されたテキストデータの検索一覧リストを表示し、

この表示された検索一覧リストの中から所望のテキストデータを指定し、

この指定されたテキストデータに含まれている前記第 1 の検索文字列の個数を前記第 1 の検索文字列の内容とともに表示し、さらに前記第 2 の検索文字列の個数を前記第 2 の検索文字列の内容とともに表示することを特徴とする情報検索方法。

【請求項 5】 前記第 1 の検索文字列の内容とこの第 1 の検索文字列が含まれている個数と、前記第 2 の検索文字列の内容とこの第 2 の検索文字列が含まれている個数とを同時に表示することを特徴とする請求項 4 記載の情報検索方法。

【請求項 6】 前記第 1 の検索文字列の内容とこの第 1 の

検索文字列が含まれている個数とを優先的に表示することを特徴とする請求項 4 記載の情報検索方法。

【請求項 7】 前記第 1 の検索文字列の内容を拡張して前記第 2 の検索文字列を作成する際に、前記第 1 の検索文字列の内容と同義語の文字列を抽出する処理を行うことを特徴とする請求項 4 記載の情報検索方法。

【請求項 8】 前記第 1 の検索文字列の内容を拡張して前記第 2 の検索文字列を作成する際に、前記第 1 の検索文字列の内容と同音異義語の文字列を抽出する処理を行うことを特徴とする請求項 4 記載の情報検索方法。

【請求項 9】 前記第 1 の検索文字列の内容を拡張して前記第 2 の検索文字列を作成する際に、文字認識の際の前記第 1 の検索文字列の内容に対して誤読傾向を示す文字列を抽出する処理を行うことを特徴とする請求項 4 記載の情報検索方法。

【請求項 10】 検索対象となる複数のテキストデータが記憶されている記憶手段と、

この記憶手段に記憶されている複数のテキストデータから所望のテキストデータを検索するために検索文字列を入力する入力手段と、

この入力手段にて入力された検索文字列を含むテキストデータを前記記憶手段から検索する検索手段と、

この検索手段にて検索されたテキストデータの検索一覧リストを表示する表示手段と、

この表示手段で表示された検索一覧リストの中から所望のテキストデータを指定する指定手段と、

この指定手段にて指定されたテキストデータに含まれている検索文字列の個数を検索文字列の内容とともに前記表示手段に表示させる制御手段と、

を具備したことを特徴とする情報検索装置。

【請求項 11】 検索対象となる複数のテキストデータが記憶されている記憶手段と、

この記憶手段に記憶されている複数のテキストデータから所望のテキストデータを検索するために第 1 の検索文字列を入力する入力手段と、

この入力手段で入力された第 1 の検索文字列を拡張して第 2 の検索文字列を作成する拡張手段と、

この入力手段で入力された第 1 の検索文字列及び前記拡張手段で拡張された第 2 の検索文字列を含むテキストデータを前記記憶手段から検索する検索手段と、

この検索手段にて検索されたテキストデータの検索一覧リストを表示する表示手段と、

この表示手段で表示された検索一覧リストの中から所望のテキストデータを指定する指定手段と、

この指定手段にて指定されたテキストデータに含まれている第 1 の検索文字列の個数を前記第 1 の検索文字列とともに前記表示手段に表示させ、さらに前記第 2 の検索文字列の個数を前記第 2 の検索文字列の内容とともに前記表示手段に表示させる表示制御手段と、

を具備したことを特徴とする情報検索装置。

10

20

30

40

50

【請求項12】 検索対象となる複数のテキストデータが記憶されている記憶手段から所望のテキストデータを検索する情報検索方法において、
前記記憶手段に記憶されている複数のテキストデータから所望のテキストデータを検索するために検索文字列を入力し、

この入力された検索文字列を含むテキストデータを前記記憶手段から検索し、

この検索されたテキストデータに含まれている検索文字列の個数を検索文字列の内容とともに表示することとを特徴とした情報検索方法。

【請求項13】 検索対象となる複数のテキストデータが記憶されている記憶手段から所望のテキストデータを検索する情報検索方法において、
前記記憶手段に記憶されている複数のテキストデータから所望のテキストデータを検索するために第1の検索文字列を入力し、

この入力された第1の検索文字列を拡張処理して第2の検索文字列を作成し、
前記第1の検索文字列及びこの第1の検索文字列を拡張した前記第2の検索文字列を含むテキストデータを前記記憶手段から検索し、

この検索されたテキストデータに含まれている前記第1の検索文字列の個数を前記第1の検索文字列の内容とともに表示し、さらに前記第2の検索文字列の個数を前記第2の検索文字列の内容とともに表示することとを特徴とする情報検索方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、検索対象情報中に含まれる文字列の検索結果を表示し、目的とする情報を容易に絞り込むことを可能とする情報検索方法および情報検索装置に関する。

【0002】

【従来の技術】 近年、情報の検索を迅速かつ容易にして、広い範囲に亘って情報を検索する方法が数多く提案されている。例えば、インターネット上で目的とする情報を検索する場合に、検索したい情報に関連すると思われる文字列を指定すると、この指定された文字列を含む情報が検索されるとともに、各々の検索結果の信頼度（指定された文字列に対する信頼度が何%というスコアで表示されるもの）を検索結果に併記して表示される方法等が知られている。但し、この検索方法においても、指定した文字列を含む情報が複数存在する場合、ディスプレイ上には、信頼度とともに複数のタイトル等が表示され、このタイトル等が指示されることによって、検索された情報の内容全てが表示される形態となっている。

【0003】

【発明が解決しようとする課題】 しかしながら従来の情報検索方法では、検索したい情報に含まれる文字列を指

定することによって、この文字列を含む情報を検索する一次検索のみしか実行していなかったため、一次検索で抽出された情報をディスプレイに表示してその内容（本文のテキスト情報）を見ながら、検索者は目的とする情報を見つけていかなければならない。そのため、検索された情報の中から目的とする情報を絞り込む作業に非常に手間がかかるという問題があった。

【0004】 本発明は、上記問題点を解決すべきなされたもので、文字列を指定することによって検索された情報（テキストデータ）の中から、目的とする情報を容易に絞り込むことを可能とする情報検索方法および情報検索装置を提供することを目的とするものである。

【0005】

【課題を解決するための手段】 上記目的を達成するために、本発明の情報検索方法においては、記憶手段に記憶されている複数のテキストデータから所望のテキストデータを検索するために検索文字列を入力し、この入力された検索文字列を含むテキストデータを記憶手段から検索し、この検索されたテキストデータの検索一覧リストを表示し、この表示された検索一覧リストの中から所望のテキストデータを指定し、この指定されたテキストデータに含まれている検索文字列の個数をこの検索文字列の内容とともに表示することとを特徴とするものである。

【0006】 すなわち、この情報検索方法では、テキストデータを検索する際に、入力した検索文字列がテキストデータの中にいくつ含まれているかを確認することで、複数のテキストデータの中から目的とするテキストデータを容易に絞り込むことができるものである。

【0007】 さらに、この発明の情報検索方法では、記憶手段に記憶されている複数のテキストデータから所望のテキストデータを検索するために第1の検索文字列を入力し、この入力された第1の検索文字列を拡張処理して第2の検索文字列を作成し、第1の検索文字列及びこの第1の検索文字列を拡張処理した第2の検索文字列を含むテキストデータを記憶手段から検索し、この検索されたテキストデータの検索一覧リストを表示し、この表示された検索一覧リストの中から所望のテキストデータを指定し、この指定されたテキストデータに含まれている第1の検索文字列の個数をこの第1の検索文字列の内容とともに表示させ、さらに第2の検索文字列の個数をこの第2の検索文字列の内容とともに表示することとを特徴とするものである。

【0008】 すなわち、この情報検索方法では、テキストデータを検索する際に、入力した検索文字列及びこの検索文字列を拡張処理された文字列がテキストデータの中にいくつ含まれているかを確認することで、複数のテキストデータの中から目的とするテキストデータを容易に絞り込むことができるものである。

【0009】

【発明の実施の形態】以下、図面を参照してこの発明の実施の形態を説明する。

（第1の実施の形態）図1は本発明の第1の実施の形態における情報検索装置の構成を示すブロック図である。

【0010】この情報検索装置においては、検索対象となるテキストデータを記憶するデータベース11、入力手段12、検索手段13、表示手段14、指定手段15、表示制御手段16とから構成されている。

【0011】データベース11は検索対象となるテキストデータからなる情報が保存されているものである。入力手段12は、検索者がデータベース11に保存されている情報の中から、所望の情報を検索するために検索文字列を入力するためのインタフェースであり、例えばキーボード等により構成されているものである。検索手段13は、検索文字列入力手段12から入力された検索文字列によりデータベース11に保存されている情報の中から、検索文字列を含むテキストデータを検索するための手段である。表示手段14は、検索手段13によって検索された情報のタイトル等のリスト一覧、検索された情報のテキストデータ等を表示するものである。指定手段15は、表示手段14に表示された検索リストの中から検索者が内容を確認したい情報を指示するためのインタフェースであり、例えばマウス等で構成されている。表示制御手段16は、指定手段15で指定された情報に対して、入力手段12にて入力された検索文字列の内容と合せて、検索手段13で検索され指定手段15で指定された情報の中に含まれている検索文字列の個数とを表示手段14に表示させるために制御する手段である。

【0012】このような構成された情報検索装置におけるテキストデータの検索処理ルーチンについて説明する。

【0013】この情報検索装置においては、図2に示す通り、検索者が検索したい情報に含まれている検索文字列を入力手段12から入力する処理を行う検索文字列入力ステップ（ST11）、入力された検索文字列に基づき、データベース11に保存されている情報の中から検索文字列を含む情報を検索する処理を行う検索ステップ（ST12）、この検索結果に基づき、検索された情報一覧リストを表示手段14に表示する検索結果表示ステップ（ST13）、表示手段14に表示された検索結果一覧リストの中から所望の情報を指定する検索文書の指定ステップ（ST14）、更に、指定された情報のテキストデータの中に検索文字列がいくつ含まれているかをカウントする検索文字列カウントステップ（ST15）並びにその結果（検索文字列がテキストデータ上にいくつ含まれているかを示す）を検索文字列とともに表示するステップ（ST16）とから成っている。

【0014】すなわち、この発明の情報検索装置においては、まず、検索文字列入力ステップ（ST11）において、検索者は入力手段12からデータベース11に保

存されているテキストデータの中から所望のデータを検索するために検索文字列（例えば、「認識」という検索文字列）を入力する。

【0015】次に、検索ステップ（ST12）にて、入力手段12から入力された検索文字列に基づき、データベース11に保存されているテキストデータの中から検索文字列を有している情報を検索し、さらに、検索結果表示ステップ（ST13）にて、表示手段14にその検索結果の一覧リストを表示するものである。

【0016】更に、検索文書指定ステップ（ST14）では、表示手段14に表示された検索結果のリストの中から検索者が情報の内容を確認したいデータを指定する。この情報検索装置においては、検索者が検索一覧リストの中から所望のものを指定した場合に、その指定されたテキストデータの内容を表示させる前に、データベース11からテキストデータを検索する際に入力された検索文字列が検索者によって指定されたテキストデータの中にいくつ含まれているかを算出し、その結果を表示させる機能を有している。すなわち、検索者によって指定されたテキストデータの中に検索の際に入力された検索文字列（「認識」という文字）がいくつ含まれているかをカウント（ST16）し、そのカウント結果を表示制御手段16により検索文字列（「認識」）とともに表示手段の所定の領域に表示させる（ST17）ものである。

【0017】図3は、その表示結果の一例を示すものであり、この表示形態では、検索一覧リストの中から所望のテキストデータの位置にマウス等の指定手段15を合せると、その位置に入力された検索文字列の用語「認識」とその検索文字列がテキストデータの中にいくつ含まれているかを示す数値「3」が表示される例を示しているものである。また、図4に示す表示形態では、検索一覧リストが表示されている領域とは別の領域に、検索文字列の用語「認識」とその検索文字列がテキストデータの中にいくつ含まれているかを示す数値「3」が表示される例を示している。

【0018】この情報検索の方法では、検索者はテキストデータ全体の内容を表示手段14で見ることにより所望のデータであるか否かを確認する前に、検索者が入力した検索文字列がテキストデータの中にいくつ含まれているかを確認することができるため、検索されたデータの全文を表示してそのデータの内容を確認する方法と比較して、文字列を指定することによって検索された情報（テキストデータ）の中から目的とする情報を容易に絞り込むことができるものである。（第2の実施の形態）図5は第2の実施の形態における情報検索装置の構成を示すブロック図である。

【0019】この情報検索装置においては、検索文字列を拡張して検索（例えば、同義語、同音異義語の文字列に拡張して検索）する場合の構成を示している。

【0020】図5に示す情報検索装置では、図1の情報検索装置と同一の構成に関しては同一符号を付している。この情報検索装置では、テキストデータを記憶するデータベース11、入力手段12、検索手段13、表示手段14、指定手段15、表示制御手段16、並びに検索文字列拡張手段21と拡張対応記憶手段22とにより構成されている。

【0021】検索文字列拡張手段21は、入力手段12から入力された検索文字列を例えば同義語や同音異議語の文字列に拡張し、また、データベース11に保存されているテキストデータが文字認識により得られたものであれば、誤読傾向に基づき文字を置換することにより文字列を拡張するといった拡張処理を行うものである。

【0022】例えば図6に示すように、検索文字列拡張手段21にて同義語による拡張処理を行う場合には、入力手段12から「コンピュータ」という検索語が入力された際に、「計算機」、「パソコン」、「PC」といった同義語に拡張処理を行うものである。

【0023】また、検索文字列拡張手段21は、同音異議語による拡張処理を行うことも可能であり、この場合には、図7に示す通り、入力手段12から例えば「対象」という検索語が入力された際に、検索者によるカナ漢字変換の誤り等の可能性を考慮して、「対照」、「対称」、「対症」等の文字の拡張を補う処理を実行するものである。

【0024】さらに、検索文字列拡張手段21は、データベース11に保存されているテキストデータが文書画像を文字認識した結果得られたものである場合には、文字が誤認識処理されている可能性を考慮して拡張処理を行うこともできる。検索文字列拡張手段21では、例えば、図8に示す通り、例えば、「ロボット」という検索語が入力された際に、誤読傾向に従い、「ロボット」という用語に対して、「ロ(漢字)ボット」、「ロボット」、「ロボット(漢字)」という検索語に拡張処理を行う。

【0025】また、拡張対応記憶手段22では、検索文字列拡張手段21にて拡張された検索文字列に対して、入力手段12にて入力された検索文字列のどの検索語に対応しているのかを記憶するためのものである。

【0026】このような構成された情報検索装置におけるデータベース11に記憶されているテキストデータの検索処理の流れについて説明する。

【0027】この情報検索装置においては、図8に示す通り、検索者が検索したい情報に含まれている検索文字列を入力手段12から入力する処理を行う検索文字列入力ステップ(ST21)、検索文字列拡張手段21により入力された検索文字列を拡張する処理を行うステップ(ST22)、入力された検索文字列に対して拡張された検索文字列を対応づけて記憶する拡張対応記憶ステップ(ST23)、拡張された検索文字列に基づき、デー

タベース11に保存されている情報の中から検索文字列を含む情報を検索するステップ(ST24)、この検索結果に基づき、検索された情報一覧リストを表示手段14上に表示する検索結果表示ステップ(ST25)、表示手段14に表示された検索結果一覧リストの中から所望の情報を指定する検索文書の指定ステップ(ST26)、更に、指定された情報のテキストデータの中に入力された検索文字列と拡張された検索文字列とが各々いくつ含まれているかをカウントする検索文字列カウントステップ(ST27)並びにその結果(入力された検索文字列と拡張された検索文字列の各々がテキストデータ上にいくつ含まれているかを示す)を検索文字列毎に表示するステップ(ST28)とから成っている。

【0028】すなわち、この情報検索の方法では、まず、検索者が入力手段12からデータベース11に保存されているテキストデータの中から所望のデータを検索するべく検索文字列(例えば、「コンピュータ」)を入力する(ST21)。

【0029】次に、入力手段12で入力された検索文字列に対して検索文字列拡張手段21にて拡張処理を実施する(ST22)。この拡張処理に関しては、上述した通り、入力手段12から入力された検索文字列を例えば同義語や同音異議語の文字列に拡張し、また、データベース11に保存されているテキストデータが文字認識により得られたものであれば、誤読傾向に基づき文字を置換することにより文字列を拡張するといった処理が行われる。すなわち、入力手段12から「コンピュータ」という検索文字列が入力された場合であれば、上述した通り、同義語の拡張処理を行い、「計算機」、「パソコン」、「PC」といった文字列を抽出する。

【0030】次に、検索文字列拡張手段21にて拡張された検索文字列(「計算機」、「パソコン」、「PC」)に対して、入力手段12にて入力された検索文字列(「コンピュータ」)とを対応づけて拡張対応記憶手段22に記憶させる(ST23)。更に、検索ステップ(ST24)にて、入力手段11にて入力された検索文字列(「コンピュータ」)と検索文字列拡張手段21にて拡張処理され拡張対応記憶手段22に記憶されている拡張された検索文字列(「計算機」、「パソコン」、「PC」)に基づき、これらの検索文字列を含むテキストデータをデータベース11から検索し、この結果を表示手段14にその検索結果の一覧リストを表示するものである(ST25)。

【0031】次に、表示手段14に表示された検索結果のリストの中から検索者が情報の内容を確認したいデータを指定する(ST26)。

【0032】この情報検索装置においては、検索者が検索一覧リストの中から所望のものを指定した場合に、その指定されたテキストデータの内容を表示させる前に、データベース11からテキストデータを検索する際に入

10

20

30

40

50

力された検索文字列、更に検索文字列拡張手段21にて拡張された検索文字列が各々テキストデータの中にいくつ含まれているかを算出し(ST27)、その結果を表示手段14に表示する(ST28)。

【0033】入力手段12から「コンピュータ」という検索文字列が入力されている場合には、検索文字列拡張手段21にて「コンピュータ」という語を「計算機」、「パソコン」、「PC」といった同義語に拡張処理されているので、図9に示す通り、拡張前の検索語(「コンピュータ」と拡張後の検索語(「計算機」、「パソコン」、「PC」)に対して、各々指定されたテキストデータに含まれている数を表示するものである。

【0034】図9に示す例では、入力手段12から入力された拡張前の検索語である「コンピュータ」という検索語が3、「コンピュータ」という検索語に対して検索文字列拡張手段21で拡張された「計算機」が2、「パソコン」が3、「PC」が0、含まれていることを表示している例を示しているものである。表示の形態としては、図9に示すように、拡張される前の検索文字列と拡張された後の検索文字列とを同時に表示する形態としてもよく、図10に示すように、まず、拡張前の検索語である「コンピュータ」と検索語が含まれている数を表示し、指定手段15等からの指示に基づき、拡張後の検索語(「計算機」、「パソコン」、「PC」)がテキストデータの中に含まれている数を表示するようにしてもよい。

【0035】さらに、この表示の形態としては、図3に示すように、検索一覧リストの中から所望のテキストデータの位置にマウス等の指定手段15を合せると、その位置に表示するようにしてもよく、また、図4に示すように、検索一覧リストが表示されている領域とは別の領域に表示するようにしてもよい。

【0036】また、図示はしていないが、入力手段12から「対象」という検索語が入力され、検索文字列拡張手段21にて「対照」、「対称」、「対症」といった同音異義語に拡張された場合には、同様に、拡張前の検索語(「対象」と拡張後の検索語(「対照」、「対称」、「対症」)に対して、各々指定されたテキストデータに含まれている検索語の数を同様に表示することもできる。更に、文字認識処理された文書画像に対しては、例えば、拡張処理される前の検索語である「ロボット」と拡張処理された後の検索語である「ロ(漢字)ボット」、「ロボット」、「ロボット(漢字)」に対しても、同様に各々指定された文書画像の中に含まれている検索語の数を表示することもできる。

【0037】すなわち、この情報検索方法においては、検索者は、検索文字列を入力することにより、検索されたテキストデータの情報の内容全体を表示手段14で見ることににより所望の情報であるか否かを確認する前に、検索文字列並びにこの検索文字列を拡張(同義語、同音

異義語、誤読傾向に基づく拡張語)した後の検索文字列が検索されたテキストデータの情報の中にいくつ含まれているかを容易に確認することが可能となる。そのため、データベース11から検索された情報の中から、検索者が必要とする情報を容易に絞り込むことができるものである。

【0038】さらに、この情報検索方法では、入力された検索文字列に対して同義語、同音異義語、誤読傾向に基づく拡張語等の文字列を抽出し、入力された検索文字列とこれらの拡張された文字列に基づいて、これらの文字列が含まれているテキストデータをデータベースから検索することができるので、検索の幅を広げることが可能となる。

【0039】尚、本発明は上述した実施の形態に限定されるものではなく、その要旨を逸脱しない範囲で種々変形実施することが可能である。

【0040】

【発明の効果】以上詳述した通り、本発明においては、記憶手段に記憶されている複数のテキストデータから所望のテキストデータを検索するために検索文字列を入力して、この検索文字列を含むテキストデータの検索を行う際に、検索されたテキストデータに含まれている検索文字列の個数を検索文字列の内容とともに表示することができるため、テキストデータを検索する際に、入力した検索文字列がテキストデータの中にいくつ含まれているかを確認することで、複数のテキストデータの中から目的とするテキストデータを容易に絞り込むことができるものである。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態に係わる情報検索装置の構成を示すブロック図である。

【図2】同装置におけるテキストデータの検索処理ルーチンを示す図である。

【図3】同装置におけるテキストデータの検索処理における表示形態の一例を示す図である。

【図4】同装置におけるテキストデータの検索処理における表示形態の一例を示す図である。

【図5】本発明の第2の実施の形態に係わる情報検索装置の構成を示すブロック図である。

【図6】同装置における検索文字列の拡張処理の一例(同義語)を示す図である。

【図7】同装置における検索文字列の拡張処理の一例(同音異義語)を示す図である。

【図8】同装置における検索文字列の拡張処理の一例(誤読傾向による処理)を示す図である。

【図9】同装置におけるテキストデータの検索処理ルーチンを示す図である。

【図10】同装置におけるテキストデータの検索処理における表示形態の一例を示す図である。

【図11】同装置におけるテキストデータの検索処理に

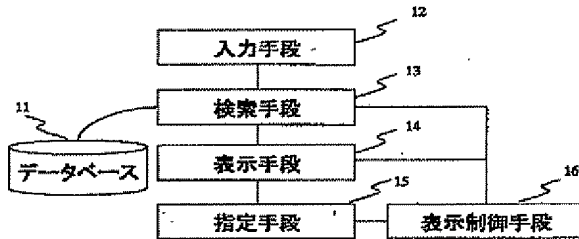
おける表示形態の一例を示す図である。

【符号の説明】

- 11 … データベース
12 … 入力手段
13 … 検索手段
14 … 表示手段
15 … 指定手段
16 … 表示制御手段

- * 14 … 表示手段
15 … 指定手段
16 … 表示制御手段
21 … 検索文字列拡張手段
* 22 … 拡張対応記憶手段

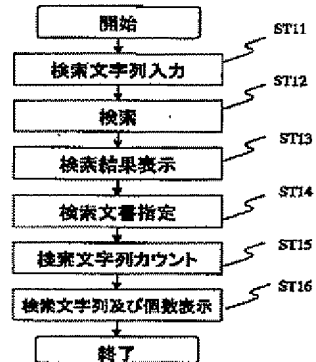
【図1】



【図3】

No	名前	237	種類	検索	認識×3	作成日
1	*****	調査	○○	*****	△△	00-9-10
2	*****	調査	○○	*****	△△	00-9-10
3	*****	調査	○○	*****	△△	00-9-9
4	*****	調査	○○	*****	△△	00-9-8
5	*****	調査	○○	*****	△△	00-9-8

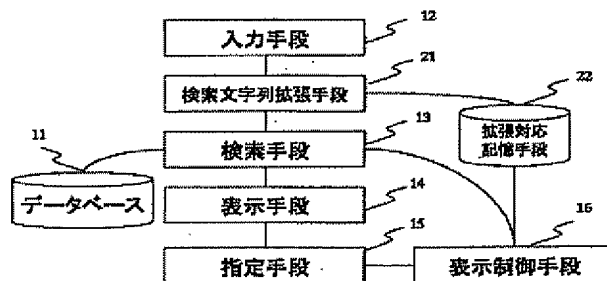
【図2】



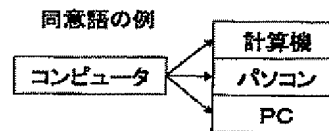
【図4】

No	名前	237	種類	検索	認識	作成日	作成日
1	*****	調査	○○	*****	△△	00-9-10	00-9-10
2	*****	調査	○○	*****	△△	00-9-10	00-9-10
3	*****	調査	○○	*****	△△	00-9-9	00-9-9
4	*****	調査	○○	*****	△△	00-9-8	00-9-8
5	*****	調査	○○	*****	△△	00-9-8	00-9-8

【図5】



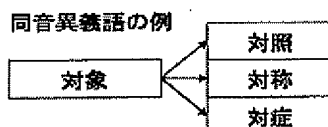
【図6】



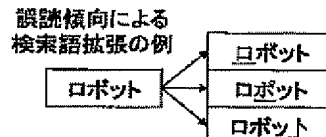
【図10】

拡張前	拡張後	HIT数
コンピュータ	計算機	2
HIT 3	パソコン	3
	PC	0

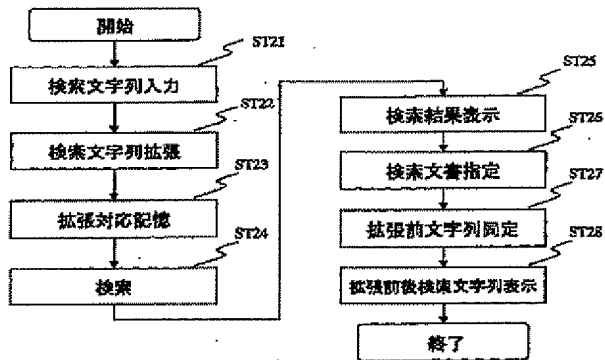
【図7】



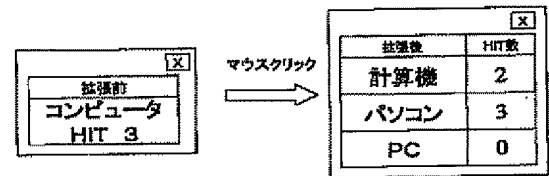
【図8】



【図9】



【図11】



フロントページの続き

(72)発明者 久保田 浩明
神奈川県川崎市幸区小向東芝町1番地 株
式会社東芝研究開発センター内

Fターム(参考) 5B075 ND03 NK02 NK35 PP02 PP03
PP13 PP22 PQ02 PQ27 PQ46
QP03